

機械学習による眼底写真に見られる 糖尿病網膜症病変の検出

Detection of diabetic retinopathy lesions seen in a fundus photograph
by machine learning

洪 時宗¹ 萩野 真一¹ 宮下 寛史¹ 飯田 修平^{2,3} 村上 聡子² 飯塚 佐知子²
小谷野 圭子^{2,3} 野村 繁之² 堀 裕士²

Sichong hong¹, Shinichi Hagino¹, Hiroshi Miyashita¹, Shuhei Iida^{2,3}, Satoko Murakami²,
Sachiko Iizuka², Keiko Koyano^{2,3}, Shigeyuki Nomura², and Hiroshi Hori²

¹株式会社システム計画研究所

¹Research Institute of Systems Planning, Inc.

²公益財団法人東京都医療保健協会 練馬総合病院

²Nerima General Hospital

³公益財団法人東京都医療保健協会 医療の質向上研究所

³Institute for Healthcare Quality Improvement, Tokyo Healthcare Foundation

Abstract: Diabetic retinopathy is one of diabetic triopathy and this ranks as the second cause of blindness for adults in Japan. Symptoms which appears in incipient diabetic retinopathy are capillary microaneurysm and small punctate bleeding, some of which can be observed by fundus photograph. In recent years, machine learning has attracted attention as a method of detecting a specific image. In this study, using supervised learning which is one of methods of machine learning, we attempt to detect the punctate bleeding seen in incipient diabetic retinopathy from fundus photograph and explore the usefulness of machine learning in fundus diagnosis.

1 概要

糖尿病の3大合併症のひとつに網膜症がある。これは血糖値が高くなることにより、細かい血管が集中している眼底部分に病変が発生する疾患である。これにより、病的な新生血管の出現や、その血管が脆いことによる出血症状の誘引が確認されている。

糖尿病網膜症の初期の所見である出血についての診断は専門医の判断が必要となっている。私たちは、専門医と機械学習の専門家を含めた議論の結果、この初期所見について、画像処理・機械学習の観点から判断のサポートをすることができると考えた。

近年インターネット上で医療関連のオープンデータが増えつつある。今回は糖尿病網膜症の眼底写真のオープンデータ [Budai, 2015] を利用し、専門医の判断を加えた上で分析を行った。

結果として病変が疑われる箇所を検出を、Precision 64.7%、Recall 38.8%で示すことができた。この結果を踏まえ、実際に臨床の現場で扱われている院内データの分析に着手する。

2 ツール作成とマーキングの依頼

機械学習を行っていく上で、まず糖尿病網膜症がどのようなものを把握する必要がある。

どの場所にどのような症状が表れるのか、専門医の立場から選ばれたデータが多数集まることで、機械学習をさせることができるからである。

専門医の意見を反映したデータを集めるために、専用の画像処理ソフトウェアを開発した。このソフトウェアは、簡単な操作による画像の表示とマーキングを可能にした。

このソフトウェアを使い、糖尿病網膜症によって生じたと疑われる病変箇所（出血・白斑・その他）のマーキングを専門医に依頼した。（図 1）またマーキングの正確なデータ（種別、座標）を回収する機能を実装した。

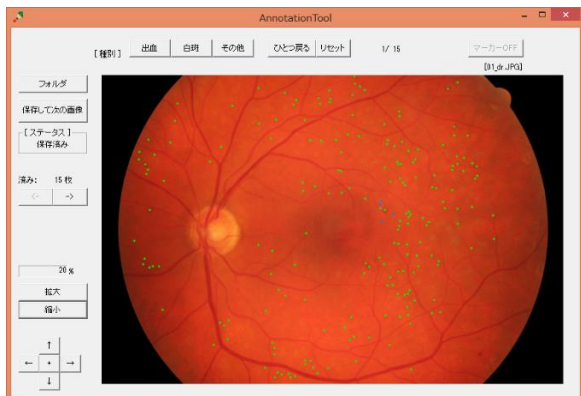


図 1 専門医によるマーキング

3 マーキングを基にした機械学習

本章では機械学習の流れについて述べる。

3.1 使用したデータ

表 1 に使用したデータの詳細を表す。

表 1 データの詳細

全データ	15 枚
内訓練データ	10 枚
内テストデータ	5 枚
画像サイズ	3504×2336 ピクセル
色情報	24 ビット

専門医によるマーキングは 15 枚の画像全てに対し依頼し、これを訓練データとして学習を行った。

3.2 前処理と学習の流れ

図 2 に学習の流れを表す。

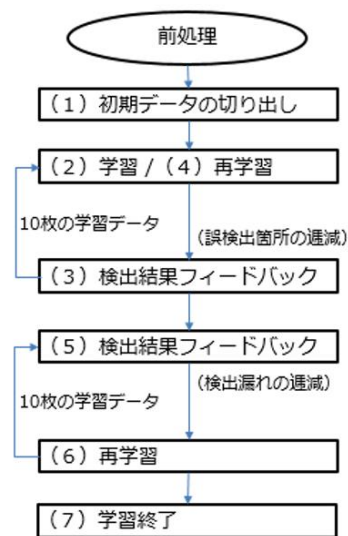


図 2 学習の流れ

3.3 前処理

専門医によるマーキングの結果を観察し、病変の特徴やデータ量の観点から特に出血箇所に着目した。

今回利用した学習手法は AdaBoost [Freund Robert E. Schapire, 1996] である。この手法は、ある決まった特徴を持つ画像を検出する際に有効な学習手法である。

網膜症の眼底写真には、出血箇所小さな円形のパターンが表れることが確認された。このパターンを学習することにより、それと同じようなパターンを持つ画像（出血箇所）を検出することができる考えた。

より有効に学習が進むように、画像に対し以下の前処理を行った。

1. 明るさの向上
2. コントラストの向上
3. G チャンネル画像への変換

これらの処理は出血箇所とそうでない周辺箇所の特徴を強調する処理である。（図 3, 図 4）

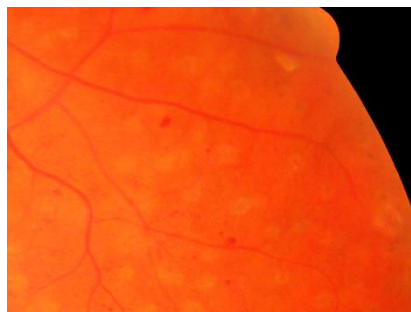


図 3 元画像の一部



図 4 前処理を行った結果

3.4 学習

AdaBoost による学習の流れについて述べる。

(1) 初期データの切り出し

専門医のマーキング結果を訓練データとし、その中で機械学習的に分かり易い病変箇所（以下 DR）と、病変でない箇所（以下 nonDR）それぞれについて画像の切り出し、ラベル付けを行った。画像の切り出しサイズは 24×24 ピクセルとした。

この際に切り出した画像に対してゆらぎ（微小位置ずらし、微小拡張）を与えることで、1 箇所の切り出しで 9 枚分の学習データを作成した。

(2) 学習

まず初期データとして用意した 108 枚（12 箇所分）の DR と、117 枚（13 箇所分）の nonDR の画像データについて AdaBoost で学習を行った。

(3) 検出結果フィードバック

学習結果を利用し、眼底写真に対して病変検出を行う。この結果、検出した病変箇所の内、誤検出された箇所（以下 nonFalseDR）にラベル付けを行う。初期の学習段階では図 5 のような血管の誤検出が多く確認された。ここでは誤検出である血管部分を nonFalseDR として学習データに加えた。

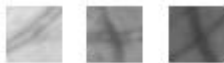


図 5 nonFalseDR 例

(4) 再学習（誤検出の通減）

nonFalseDR を加え再度学習を行う。再学習は血管が検出されなくなる程度まで繰り返した。

(5) 検出結果フィードバック

血管が検出されなくなるのに伴い、病変の検出漏れが増加しはじめる。再度訓練データを基に、検出漏れが起きている病変に対して図 6 のような DR のラベル付けを行う。

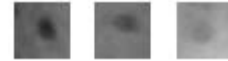


図 6 DR 例

(6) 再学習（検出漏れの通減）

DR を学習データに加えた上で、再度学習を行う。適宜血管が検出された場合は nonFalseDR のラベル付けを行う。

(7) 学習終了

学習が十分に進むと、血管の検出を減らす操作で検出漏れが増え、検出漏れを減らす操作で血管の誤検出が増える兆候が表れる。この段階まで学習が進むと誤検出、検出漏れが共に目立たなくなる。

検出結果が専門医のマーキングに比較的近い状態になった時点で学習を終了した。

4 テストデータでの検出

学習結果を基に、テスト用データ 5 枚に対して検出を行った。

学習による病変検出箇所について中心座標を CSV ファイルで出力し、正解データと比較して座標が 25 ピクセル未満であった際は True、それ以外は False とした。正解データとして用いたのは、専門医によるマーキングのデータである。

True データを TruePositive、誤検出された箇所を FalsePositive、検出漏れを FalseNegative とし、結果を表 2 に示す。

表 2 正解データと検出結果

正解データ	340
TruePositive(TP)	132
FalsePositive(FP)	72
FalseNegative(FN)	208

検出精度に対する一般的な指標として正確性（適合率）を表す Precision と検出の網羅性（再現率）を表す Recall がある。それぞれ以下の結果となった。

$$\begin{aligned} \text{Precision} &= \frac{TP}{TP + FP} \\ &= 0.647 \end{aligned}$$

$$\begin{aligned} \text{Recall} &= \frac{TP}{TP + FN} \\ &= 0.388 \end{aligned}$$

5 考察

機械学習の結果、正解データに対して Precision が 64.7%、Recall が 38.8%となった。

今回のアプローチでは、院内データの分析に展開するために、専門医に依頼したマーキングについて病変の進行度合いや大きさなどの、細かな制約を設けなかった。そのため、病変の疑いがあるとされたものの、訓練データとして扱いつらいデータが多数確認された。

眼底写真の種類によっては Recall に大きな差があり、18.8%と低くとどまったものから、81.3%と高い結果を示すものもあった。このことから検出する対象となる病変の特徴についての基準を限定させることで、病変と思われる箇所網羅性についてより高い結果を示すことができると考えられる。

6 今後の課題

今後は練馬総合病院の実患者データの分析に着手する。

院内データは、オープンデータと比較してデータサイズが小さな画像であるため、如何に効果的な前処理を行えるかが鍵となる。

一方、学習に利用できる画像枚数が多く、今回とは別の分析アプローチも取れると考えている。例えば、病変が疑われる箇所の判定だけでなく、1枚の眼底写真に対して、健康か又は病気が疑われるか、といった判定を、院内で蓄積されている電子カルテに紐付くラベル付けに対応させて学習させるといった方法である。

参考文献

- [1] 中野哲平: 医療画像における局所特徴量の検知, The 28th Annual Conference of the Japanese Society for Artificial Intelligence, 2014:301-10in.
- [2] 「High-Resolution Fundus (HRF) Image Database」
<<https://www5.cs.fau.de/research/data/fundus-images/>>
(2015/09/15 アクセス)
- [3] Yoav Freund and Robert E. Schapire. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. Journal of Computer and System Sciences, 55(1):119--139, August 1997.